IN THE SPECIFICATION

Please replace the paragraph beginning on page 17, line 10 and ending on page 18, line 4 with the following amended paragraph:

Next a test is made in step 406 whether the SOLA operation will be applied to compress (i.e. speed up) or expand (i.e. slow down) two adjacent windows of speech samples. In the case where the speech is to be expanded the process continues in step [[408]] 418. As shown in FIG. 6, the speech sample in A2 oldwin 502 [[are]] is copied and inserted in between oldwin A2 and [[B2]] B1 newin [[508]] 504 as shown as A2 oldwin 602. Next the pointer for oldwin 510 is adjusted to point to the beginning of A2 oldwin 602. Therefore a duplicate 602 of window 502 now exists in the buffer. Next, the process continues in step 410 for both compression and expansion of the speech window (after a copy of the oldwin was inserted in step 408 for expansion). A cross correlation (oldwin, newin) is performed to determine the position of maximum correlation. In the case where an expansion is being performed the oldwin is A2 oldwin 602 as shown in FIG. 6 and for a compression the oldwin is A2 oldwin 502 as shown in FIG. 5. Next, in step 412, the newin pointer 512 is set to the position of maximum correlation. Finally, ~~Next~~ in step 414 a real-time SOLA (Synchronized OverLap and Add) operation is performed where the data is written, using a write pointer, into the audio output buffer at a position of maximum correlation and the process ends at step 416. The output of the SOLA operation in place on the circular buffer is shown for compression in graph 702 and for expansion in graph 704 of FIG. 7. Notice the SOLA region AB 712 looks a lot like the SOLA region AB 722. This is because of the blending of A2 and B2 in both cases. Specifically for compression, as illustrated in graph 702, the oldwin A2 and newin B1 are now combined into one region AB 712. For expansion, as illustrated in graph 704, oldwin A2 was copied prior to being blended. This provides an over view of the operation. Different rates are achieved by varying the

size the region of overlap to which the SOLA is performed. Stated differently, for two times, compression or expansion, the region A1+A2 is used instead of simply A2, and the region B1+B2 is used instead of simply B2. Next, more specifics of each step in the flow chart of FIG. 4 are now described.

Please replace the paragraph on page 18, lines 5-23 with the following amended paragraph:

In one embodiment, the present invention uses FOUR rate adjustments used in the audio playback speed: Very slow (~1.7x), slow (~1.4x), fast (~0.8x) and very fast (~0.6x), where x describes the multiplicative change in time. So very slow means it plays it back 1.7 times as slow. These numbers are approximate rate changes, since the procedure is dependent on the ~~speakers~~ speaker's pitch. All four modes utilize the SOLA method. It is important to note that other number and rates of playback are within the true scope and spirit of the present invention. The different modes are selected by the state algorithm configuration which is one of two states: expansion or compression, and of one of two levels: half or full. In full compression, the SOLA blending is performed on every entrant (new) frame. In half compression, the SOLA blending is performed on every other frame, and requires only a simple flag. The expansion mode is essentially the same as compression and only requires a frame duplication before the SOLA method. In full expansion, every frame is duplicated before the SOLA method. In half expansion, the frame replication and SOLA method are performed on every other frame. The half rate selection is a simple integration effort since it only requires a pointer location update. A flag is not required for half expansion. FIG. [[6]] 8 is a state diagram [[600]] illustrating the algorithm configurations for each of the four rate selections, which are discussed in further detailed in the section entitled "Detailed Overview Of SOLA Speech Time Compression" below.

Please replace the paragraph on page 20, lines 13-29 with the following amended paragraph:

STEP 4: FIG. 12, is a graph ~~1100~~ of a composite speech signal for which windows oldwin A2 502 and newin B1 [[502]] 504 have been blended into an AB subframe 1112 for speech compression, according to the present invention. Leave speech between beginning of oldwin pointer 510 and runindex1on outbound audio buffer. This region does not contain the current processed SOLA region. The SOLA region, which was the overlapped speech region, is processed directly on the buffer. The remainder of newin B1 504 must be shifted towards oldwin A2 502 to append the unmodified data. Runindex1 will now specify the beginning of oldwin 602 for the next SOLA cycle. Steps 1 to 4 are repeated and performed for each new frame of speech placed on the audio buffer. This is how time compressed speech is generated. These steps also allow the dynamic rate adjustment for SOLA as speech is being played out since the compression is adjustable on a frame-by-frame basis. This is covered in the next section. NOTE: SOLA operates on data in the outbound voice buffer. Both frames (oldwin A2 502, newin B1 504) are sequential and adjacent in the outbound buffer. The SOLA processing is performed in place in the buffer and the OB_Voice_Wptr (as described below) must be updated accordingly. The SOLA support routines in the next sections perform the ob_voice_buf_wptr updates.